MEME15203 Statistical Inference Marking Guide

Assignment 5

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME15203
Course:	MAC	Unit Title:	Statistical Inference
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	January 2022		
Due by:			

Q1. Let X have probability density function

$$f(x) = \begin{cases} \frac{\Gamma(7)x^4(\theta-x)}{\Gamma(5)\theta^6}, & 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}$$

Show that $\frac{X}{\theta}$ is a pivotal quantity and use this pivotal quantity to find a 92% lower confidence limit for θ.

(20 marks)

Ans.

Let $U = \frac{X}{\theta}$, then

$$\begin{aligned} F_U(u) &= P(U \leq u) \\ &= P(\tfrac{X}{\theta} \leq u) \\ &= P(X \leq u\theta) \\ &= \int_0^{u\theta} \frac{\Gamma(7)x^4(\theta-x)}{\Gamma(5)\theta^6} dx \\ &= \frac{\Gamma(7)}{\Gamma(5)\theta^6} \int_0^{u\theta} [\theta x^4 - x^5] dx \\ &= \frac{\Gamma(7)}{\Gamma(5)\theta^6} \left[\frac{\theta x^5}{5} - \frac{x^6}{6}\right]_0^{u\theta} \\ &= \frac{\Gamma(7)}{\Gamma(5)\theta^6} \left[\frac{u^5\theta^6}{5} - \frac{(u\theta)^6}{6}\right] \\ &= \frac{\Gamma(7)}{\Gamma(5)} \left[\frac{u^5}{5} - \frac{u^6}{6}\right] \\ &= 6u^5 - 5u^6 \end{aligned}$$

which is free of θ, $\therefore U = \frac{X}{\theta}$ is a pivotal quantity.

Set $P(U \leq a) = 0.92$ so that $P(\frac{X}{a} < \theta) = 0.92$

[Note; I have make a mistake by using $P(U < b)$ for upper bound, which is not correct, please rectify yourself.]

$6a^5 - 5a^6 = 0.92$

$0.92 - 6a^5 + 5a^6 = 0$

$b = polyroot(c(0.92, 0, 0, 0, 0, -6, 5)) = 0.918431$

So a 92% lower confidence limit for θ is $\frac{x}{0.918431}$.

Q2. Consider independent random samples from two normal distributions, $X_i \sim N(0, a_1)$ and $Y_j \sim N(0, a_2)$; $i = \ldots, 30$, $j = 1, \ldots, 30$. Derive a $100(1-\alpha)\%$ confidence interval for $\frac{a_1}{a_2}$ based on sufficient statistics.

(20 marks)

Ans.

$f(x_1, \ldots, x_{30}; a_1) = (2\pi)^{-30/2} a_1^{30} e^{-\sum_{i=1}^{30} x_i^2/2a_1} = g(s, a_1)h(x_1, \ldots, x_{30})$

where $s_1 = \sum_{i=1}^{30} x_i^2$, by factorization theorem, s_1 is a sufficient statistic for a_1.

Similarly, $s_2 = \sum_{j=1}^{30} y_j^2$ is a sufficient statistic for a_2.

Note that $\frac{\sum_{i=1}^{30} x_i^2}{a_1^2} \sim \chi^2(30)$ and $\frac{\sum_{j=1}^{30} y_j^2}{a_2} \sim \chi^2(30)$, and

$\frac{\frac{\sum_{j=1}^{30} y_j^2}{30a_2}}{\frac{\sum_{i=1}^{30} x_i^2}{30a_1}} = \frac{a_1}{a_2}\frac{30}{30}\frac{\sum_{i=1}^{30} y_i^2}{\sum_{j=1}^{30} x_j^2} \sim F(30, 30)$

$P\left(F_{\alpha/2}(30,30) < \frac{a_1}{a_2}\frac{30}{30}\frac{\sum_{i=1}^{30} y_i^2}{\sum_{j=1}^{30} x_j^2} < F_{1-\alpha/2}(30,30)\right) = 1-\alpha$

$P\left(\frac{30}{30}\frac{\sum_{j=1}^{30} x_j^2}{\sum_{i=1}^{30} y_i^2} < F_{\alpha/2}(30,30) < \frac{a_1}{a_2} < \frac{30}{30}\frac{\sum_{j=1}^{30} x_j^2}{\sum_{i=1}^{30} y_i^2} F_{1-\alpha/2}(30,30)\right) = 1-\alpha$

Thus $100(1-\alpha)\%$ confidence interval for $\frac{a_1}{a_2}$ is

$$\left(\frac{30}{30}\frac{\sum_{j=1}^{30} x_j^2}{\sum_{i=1}^{30} y_i^2} F_{\alpha/2}(30,30), \frac{30}{30}\frac{\sum_{j=1}^{30} x_j^2}{\sum_{i=1}^{30} y_i^2} F_{1-\alpha/2}(30,30)\right)$$

Q3. Consider independent random samples from two exponential distributions, $X_i \sim EXP(\mu)$ and $Y_j \sim EXP(\lambda)$; $i = 1, \ldots, 30$, $j = 1, \ldots, 30$.

(a) Find the distribution of $(\lambda/\mu)(\bar{X}/\bar{Y})$.

(b) Derive a $100(1-\alpha)\%$ confidence for λ/μ.

(20 marks)

Ans.

$\frac{60\bar{X}}{\mu} \sim \chi^2(60)$, $\frac{60\bar{Y}}{\lambda} \sim \chi^2(60)$

Note that if $X \sim GAM(\alpha, \theta)$, then $2X/\theta \sim \chi^2(2\alpha)$, in example 2 of Chapter7, I make a mistake by stating that if $\sum X_i = n\bar{X} \sim GAM(4n_1, \beta_1)$, then $2n_1\bar{X}/\beta_1 \sim \chi^2(2n_1)$ which is not correct, it should be $\chi^2(8n_1)$

$\frac{\frac{60\bar{X}}{60\mu}}{\frac{60\bar{Y}}{60\lambda}} = \frac{\lambda\bar{X}}{\mu\bar{Y}} \sim F(60, 60)$.

$P\left(F_{\alpha/2}(60,60) \le \frac{\lambda\bar{X}}{\mu\bar{Y}} \le F_{1-\alpha/2}(60,60)\right) = 1-\alpha$

$P\left(\frac{\bar{Y}}{\bar{X}}F_{\alpha/2}(60,60) \le \frac{\lambda}{\mu} \le \frac{\bar{Y}}{\bar{X}}F_{1-\alpha/2}(60,60)\right) = 1-\alpha$

Thus, a $100(1-\alpha)\%$ confidence for λ/μ is

$$\left(\frac{\bar{y}}{\bar{x}}F_{\alpha/2}(60,60), \frac{\bar{y}}{\bar{x}}F_{1-\alpha/2}(60,60)\right).$$

Q4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pdf

$$f(x|\lambda) = \frac{\lambda^5}{\Gamma(5)} x^4 e^{-\lambda x}, x > 0, \text{ zero othewise,}$$

the prior density of λ is

$$\pi(\lambda) = \frac{\mu^3}{\Gamma(3)} \lambda^2 e^{-\mu\lambda}, \lambda > 0, \text{ zero othewise,}$$

where μ is known. Derive a $100(1-\alpha)\%$ equal probabilty Bayesian confidence interval for λ in terms of χ^2 random variable.

(20 marks)

Ans.

$f(x_i|\lambda) = \frac{\lambda^5}{\Gamma(5)} x^4 e^{-\lambda x_i}, x_i > 0$

$\pi(\lambda) = \frac{\mu^3}{\Gamma(3)} \lambda^2 e^{-\mu\lambda}, \lambda > 0$

$\pi(\lambda|\mathbf{x}) = k\lambda^{5n} e^{-\lambda n\bar{x}} \lambda^2 e^{-\mu\lambda} = k\lambda^{5n+2} e^{-\lambda(\mu+n\bar{x})}$

$\Rightarrow \Lambda|\mathbf{x} \sim Gamma(5n+3, (\mu+n\bar{x})^{-1})$

Note that $2(\mu+n\bar{x})\Lambda \sim Gamma(5n+3, 2) \sim \chi^2(2(5n+3))$

$P(\chi^2_{\alpha/2}(\nu) \leq 2(\mu+n\bar{x})\lambda) \leq \chi^2_{1-\alpha/2}(\nu)) = 1-\alpha$

$P\left(\frac{\chi^2_{\alpha/2}(\nu)}{2(\mu+n\bar{x})} \leq \lambda \leq \frac{\chi^2_{1-\alpha/2}(\nu)}{2(\mu+n\bar{x})}\right) = 1-\alpha$

where $\nu = 2(5n+3)$. Thus a $100(1-\alpha)\%$ equal probabilty Bayesian confidence interval for λ is

$$\left(\frac{\chi^2_{\alpha/2}(\nu)}{2(\mu+n\bar{x})}, \frac{\chi^2_{1-\alpha/2}(\nu)}{2(\mu+n\bar{x})}\right).$$

Q5. Losses follow a gamma distribution with $\alpha = 3$ and θ unknown. The prior distribution of θ has density function $\pi(\theta) = \frac{1}{\theta}$, Five losses are observed:

$$[629.6, 178.7, 189.4, 127.0, 665.1].$$

Determine the 95% HPD credible interval for θ.

(10 marks)

Ans.

$f(x_i|\theta) = \frac{1}{\Gamma(3)\theta^3} x_i^2 e^{-x_i/\theta}$

$\pi(\theta|\mathbf{x}) = k\theta^{-5(3)} e^{-\sum x_i/\theta} \theta^{-1} = k\theta^{-16} e^{-1789.8/\theta}$

$\Rightarrow \Theta|\mathbf{x} \sim Invgamma(\alpha' = 15, \theta = 1789.8)$

$a = qinvgamma(0.025, 15, 1789.8) = 76.1954$

$b = qinvgamma(0.975, 15, 1789.8) = 213.1885$

Using R to solve the simultaneous equation, the 95% HPD credible interval for θ is $(69.1062, 198.5092$.

```
R-Codes:
#First install nleqslv and invgamma packages by typing:
install.packages("nleqslv") and
install.packages("invgamma")
library(invgamma)
f = function(x){
y = numeric(2)
y[1] = pinvgamma(x[2],15,1789.8) - pinvgamma(x[1],15,1789.8) - 0.95
y[2] = dinvgamma(x[1],15,1789.8) - dinvgamma(x[2],15,1789.8)
y
}
library(nleqslv)
xstart = c(76.1954,213.1885)
nleqslv(xstart, f, control=list(btol=.01),
method="Newton")
```

Q6. You are given:

$$f(x|\theta) = \begin{cases} (\theta+1)x^\theta & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\pi(\theta) = \begin{cases} \frac{1}{\theta+1} & \text{for } \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

Suppose that a single observation takes the value $x = 0.33$. Find the upper bound of the 98% HPD credible region for θ.

(10 marks)

Ans.

$\pi(\theta|x = 0.33) = k(0.33^\theta) = ke^{\theta \ln(0.33)}, \theta > 0$
$k \int_0^\infty e^{\ln(0.33)\theta} d\theta = 1$
$k \left[\frac{e^{\ln(0.33)\theta}}{\ln(0.33)} \right]_0^\infty = 1$
$k = -\ln(0.33)$
$\pi(\theta|x = 0.33) = -\ln(0.33)e^{\theta \ln(0.33)}, \theta > 0$
Since $\pi(\theta|x = 0.33)$ is a increasing function of θ,the credible set will be of the form $(0, u)$
$\int_0^u -\ln(0.33)e^{\theta \ln(0.33)} d\theta = 0.98$
$\left[-e^{\ln(0.33)\theta} \right]_0^u = 0.98$

$$1 - e^{\ln(0.33)u} = 0.98$$
$$u = \frac{\ln(0.02)}{\ln(0.33)} = \boxed{3.528596}$$